ADF COMPANION DOCUMENT C: Description of the Data Intrusion Simulation (DIS) Method.

1 Introduction

The concept behind the DIS method derived from concerns expressed by Elliot (1996) regarding the need to examine statistical disclosure risk from the viewpoint of the data intruder (intruder-centrically) rather than from that of the data themselves (data-centrically). A rational intruder would be indifferent to questions such as, for example, whether a record was sample or population unique, because s/he will know such attributions of status are unreliable and more importantly because s/he will have more pragmatic concerns, such as whether her/his actual matches are correct. The DIS method simulates the intruder perspective by focusing on the probability of a unique match being correct. The basic assumption is that the intruder has some information about a population unit and uses that information to attempt to find the record for that individual in a microdata file (which is a sample of the relevant population). If there is only one record in the dataset which corresponds to the information that the intruder has that is called a unique match. If that record is the correct record for that population unit that is called a correct match. These basic elements form the headline statistic of a DIS analysis; the probability of a correct match given a unique match: pr(cm|um).

The basic principle of the DIS method is to remove records from the target microdata file and then re-sample them according to the original sampling fraction (the proportion of the population that are in the sample). This creates two files, a new, slightly truncated, target file and a file of the removed records which can then be matched against the target file. The method has two computational forms, the *special form*, where the sampling is actually done, and the *general form*, where the sampling is not actually performed, but its effect is derived using the equivalence class structure and sampling fraction.

2 The special method

The special DIS method uses a similar technique to Briggs (1992).

1. Set counters U and C to zero.

2. Take a sample microdata file (A) with sampling fraction S.
3. Remove a random record (R) from A, to make a new file (A').
4. Generate a random number (N) between 0 and 1. If N<=S then copy back R into A' with each record having a probability of being copied back equal to S.
5. The result of this procedure is that B will now represent an arbitrary population unit whose probability of being in A' is equal to the original sampling fraction.
6. Match fragment against A'. If R matches a single record in S' then add record 1 to U if the match is correct add 1 to C.
7. Iterate through stages ii-v until C/U stabilises.

3 The general method

A more general method can be derived from the above procedure. Imagine that the removed fragment (B) is just a single record. Clearly there are six possible outcomes depending on whether the record is resampled or not and whether it was a unique, in a pair, or in a larger equivalence class.

Table 1: Possible per record outcomes from the DIS general method

record is:	*Copied back*	*Not copied back*
sample unique	**correct unique match**	non-match
one of a sample pair	multiple match including correct	**false unique match**
one of a larger equivalence class	multiple match including correct	false multiple match

Given this, one can derive the estimated probability of a correct match given a unique match from:

$$\frac{U \times \Pi}{U \times \Pi + P \times (1 - \Pi)}$$

Where U is the number of sample uniques, P is the number of records in pairs and Π is the sampling fraction.

For full statistical proof of the above theory see Skinner and Elliot (2002). For a description of an empirical study that demonstrates that the method works see Elliot (2000). For an elaboration using the special method for post-perturbation disclosure

risk assessment see Elliot (2001). For an extension which takes account of general misclassification errors see Elamir and Skinner (2006).